

# Book Reviews

Edward Swanson

**Essential Thesaurus Construction.** By Vanda Broughton. London: Facet Pub., 2006. 296p. \$65 paper (ISBN 978-1-85604-656-0/1-85604-565-X).

**The Thesaurus: Review, Renaissance, and Revision.** Eds. Sandra K. Roe and Alan R. Thomas. Binghamton, N.Y.: Haworth Infor. Pr., 2004. 209p. \$39.95 cloth (ISBN 07890-1978-7); \$19.95 paper (ISBN 07890-1979-5). Published simultaneously as *Cataloging & Classification Quarterly*, 37, nos. 3/4.

As stated in the subtitle of Sandra Roe and Alan Thomas's collection of essays, the thesaurus may indeed be experiencing a renaissance in our digital era. Many information professionals continue to insist that access to electronic full-text is the only important precondition for resource discovery on the Web. At the same time, others, including professionals outside librarianship proper, have come to understand that the mediation provided by expertly developed and maintained controlled vocabularies is a basic requirement for service to populations of all sorts (including impatient undergraduates), and not an unwarranted interference with the end-user's autonomy. The two volumes under review, together, address this topic in complementary fashion.

Vanda Broughton, lecturer in library and information studies at the School of Library, Archive and Information Studies, University College London, has written a valuable manual for students and practitioners. *Essential Thesaurus Construction* is concerned with both the principles and practice of thesaurus construction, "with rather more emphasis than is usual on the latter" (1). She has at the same time provided a practical and detailed introduction to taxonomy construction, outlining the basic methods for building a subject vocabulary. This is useful not only for the student, but also for the information professional who must create a thesaurus in-house.

Broughton's book, while emphasizing practical application in real-world situations, does not slight theoretical issues. Instead, chapters with either theoretical or practical emphases are integrated in a single logical sequence. The opening chapters discuss fundamentals, such as the nature of a thesaurus and how it is distinguished from other subject access tools, uses of thesauri and their advantages, types of thesauri, and the different displays typically provided. Practical steps are then described, beginning with five chapters on aspects of vocabulary selection and simple vocabulary control, a process that results in the needed raw material from which a thesaurus is constructed. Chapters

on thesaural relationships, facets and arrays, hierarchies, and the complex issues surrounding compound subjects and citation order, demonstrate the development of the thesaurus from a mass of unstructured terminology to a logically developed system. The final chapters concern conversion of a classified arrangement to alphabetical format, creation of thesaurus records, and methods for maintenance and updating. The entire process is illustrated at every stage through the actual development of a thesaurus on animal welfare, beginning with basic vocabulary sources such as scholarly journal articles and Web resources, through to the presentation of fully structured entries. It becomes clear that, while the thesaurus and taxonomy construction may not be for the timorous, there is a well-marked path to success for the determined and careful beginner.

Clear expositions of often difficult concepts enhance the text. These include the thesaurus as indexing tool versus organizational or navigational tool (34) and the relationship of polyhierarchy to notation (179). The discussion of compound terms, and the circumstances under which they should be factored into simpler terms (beginning on page 90), prepares the reader for the complex question of whether to provide Broader/Narrower Term (BT/NT) relationships, or Related Term (RT) relationships, to thesaurus terms which are retained as compounds (180).

Valuable pedagogical features complement Broughton's lucid prose. Glossary terms, when first appearing in the text, are in bold face. The glossary itself is written for non-specialists, with an emphasis on "helpful explanations . . . rather than precise technical definitions" (208). Most chapters feature several summaries, allowing for a quick review of new material. Exercises, with answers, are introduced beginning with chapter 11, concerning term extraction from document titles. Following the glossary, six appendices allow the motivated reader to examine the development of the sample thesaurus in detail.

There were very few errors noted in this well-produced volume, which is very convenient in format and easily lies flat. On page 219 there is the phrase, "one of the Ranganathan's fundamental categories." The reference on page 248 to the animal product "fu" is baffling, until one realizes that "fur" is meant. On page 261, the class notation GAP seems out of order and probably should have been GP instead. These could be easily handled in an updated printing or revised edition. Finally, one of the incidental pleasures of this sort of book is how it expresses, even through the examples given, the vastness and complexity of the spheres of knowledge

and human activity. (See, for example, the terms *screaming hairy armadillo*, *clog dancing*, *coastal erosion*, and *Biblical hermeneutics*, found on pages 88–89.)

Roe and Thomas's *The Thesaurus: Review, Renaissance, and Revision*, published two years prior to Broughton's work, nevertheless benefits by being read in conjunction with the latter. (Disclosure: the reviewer is presently a member of the *Cataloging & Classification Quarterly* editorial board, but was not at the time of this volume's publication.) The editors describe three motivations for producing this volume: "to acquaint or remind the Library and Information Science (LIS) community of the history of the development of the thesaurus and the standards for thesaurus construction . . . to provide bibliographies and tutorials from which any reader can become more grounded in her or his understanding of thesaurus construction, use, and evaluation . . . [and] to address topics related to thesauri but that are unique to the current digital environment" (1–2). As with all such volumes of collected essays, it is almost inevitable that readers will find different contributions to be of greater or lesser value or interest, but all three of the editors' motivations are successfully addressed to different extents.

"The Thesaurus: A Historical Viewpoint, with a Look to the Future," by Jean Aitchison and Stella Dextre Clarke, appropriately opens the volume. Aitchison, who compiled the pioneer work *Thesurofacet* in 1969, and is responsible for many faceted thesauri, is coauthor of *Thesauri Construction and Use*, now in its fourth edition and frequently cited by Broughton.<sup>1</sup> Following a historical review of the development of the thesaurus as a type of tool, Aitchison and Dextre point out the "two major trends in thesaurus development today . . . for adaptations that will make a controlled vocabulary much quicker, easier, and more intuitive to use . . . [and for] interoperability" (14). The cry for more intuitive tools, however, runs into the contradiction between more loosely defined popular taxonomies on the one hand, and the need for "much more precisely defined term relationships" (16) for Semantic Web use on the other.

The next three articles address the intention to provide background readings and exercises for the novice. Alan R. Thomas's "Teach Yourself Thesaurus: Exercises, Readings, Resources" provides readings grouped in topical categories, to serve as a basis for "self-instruction in thesaurus-making" (24). Specific exercises are not provided, but examples published elsewhere are pointed out. Marianne Lykke Nielsen's "Thesaurus Construction: Key Issues and Selected Readings" is a more conventional bibliographic essay, with readings organized generally around the stages of thesaurus construction. "A Practical Exercise in Building a Thesaurus," by James R. Shearer, aims to cover in twenty-two pages what *Essential Thesaurus Construction* covers in nearly three hundred. It is difficult to picture how this extremely condensed exposition could really be of use to

the beginner. However, Shearer's essay, along with those by Thomas and Nielsen, could be of value as a supplement to Broughton's work.

"Thesaurus Consultancy" by Leonard Will and "Thesaurus Evaluation" by Leslie Ann Owens and Pauline Atherton Cochrane describe professional services of value in the construction, testing, and revision stages of thesaurus development. The breadth of the latter essay is worthy of note. Owens and Cochrane discuss multiple approaches to evaluation—the "comparative, observational, formative, and structural methods"—and their applications in myriad contexts, "online or printed [thesauri], machine or human-generated, stand-alone or integrated, monolingual or bilingual, standards-compliant or not" (99). End-user research is represented in this volume by Jane Greenberg's "User Comprehension and Searching with Information Retrieval Thesauri." Business school graduate students searching ABI/Inform were studied regarding their past experience with use of thesauri in online searching, their desire to use thesauri once introduced to them, and their preferred "processing methods" (112) when working with thesauri. One outcome is the conclusion that users prefer "either interactive or a combination of automatic and interactive thesaurus processing to completely automatic processing" (116).

Of the remaining contributions, Eric H. Johnson's "Distributed Thesaurus Web Services" and Patrice Landry's "Multilingual Subject Access: The Linking Approach of MACS" are of great interest. Johnson's forward-looking essay tackles the challenge of making the online subject thesauri much more useful for searching by the general Web-using public. He describes the Thesaurio-Web, "a proposed network of thesaurus access and navigation services" (121), using a distinctly developed XML-based markup language and user interface. The idea that "you can search the Web more easily and effectively using specialized search applications, only using the Web browser to fetch and display the actual Web documents" (127) is intriguing. Finally, Landry's description of the European MACS (Multilingual Access to Subjects) project, while not primarily concerned with thesaurus development per se, is valuable in that it raises the issue of interoperability, not only among languages, but more broadly among disciplines and controlled vocabularies of different levels of granularity. Taken together, Johnson and Landry remind us of the work that still needs to be done to provide access to the digital universe, by means that are as intelligent as they are intelligible.—David Miller, *dmiller@curry.edu*, Curry College, Milton, Mass.

#### Reference

1. Jean Aitchison, Alan Gilchrist, and David Bawden, *Thesaurus Construction and Use: A Practical Manual*, 4th ed. (London: Aslib, 2000).